Adaptive steered molecular dynamics: Validation of the selection criterion and benchmarking energetics in vacuum

Gungor Ozer, Stephen Quirk, and Rigoberto Hernandez

Citation: J. Chem. Phys. **136**, 215104 (2012); doi: 10.1063/1.4725183
View online: http://dx.doi.org/10.1063/1.4725183
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v136/i21
Published by the American Institute of Physics.

Adaptive steered molecular dynamics: Validation of the selection criterion and benchmarking energetics in vacuum

Gungor Ozer,[1] Stephen Quirk,[2] and Rigoberto Hernandez[1,a)]
[1]*Center for Computational and Molecular Science and Technology, School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA*
[2]*Kimberly-Clark Corporation, Atlanta, Georgia 30076-2199, USA*

(Received 17 January 2012; accepted 18 May 2012; published online 7 June 2012)

The potential of mean force (PMF) for stretching decaalanine in vacuum was determined earlier by Park and Schulten [J. Chem. Phys. **120**, 5946 (2004)] in a landmark article demonstrating the efficacy of combining steered molecular dynamics and Jarzynski's nonequilibrium relation. In this study, the recently developed adaptive steered molecular dynamics (ASMD) algorithm [G. Ozer, E. Valeev, S. Quirk, and R. Hernandez, J. Chem. Theory Comput. **6**, 3026 (2010)] is used to reproduce the PMF of the unraveling of decaalanine in vacuum by averaging over fewer nonequilibrium trajectories. The efficiency and accuracy of the method are demonstrated through the agreement with the earlier work by Park and Schulten, a series of convergence checks compared to alternate SMD pulling strategies, and an analytical proof. The nonequilibrium trajectories obtained through ASMD have also been used to analyze the intrapeptide hydrogen bonds along the stretching coordinate. As the decaalanine helix is stretched, the initially stabilized $i \rightarrow i + 4$ contacts (α-helix) is replaced by $i \rightarrow i + 3$ contacts (3_{10}-helix). No significant formation of $i \rightarrow i + 5$ hydrogen bonds (π-helix) is observed.
 [http://dx.doi.org/10.1063/1.4725183]

I. INTRODUCTION

The stability and function of proteins hinge on the relative free energies of contacts within a chain and contacts between its residues and the solvent. Several computational schemes are being developed in order to better characterize both of these types of contact free energies for systems described at the molecular scale—viz. so-called physics-based models. This includes, but is not limited to, replica exchange MD,[1] adaptive biasing force MD,[2] and free energy perturbation MD.[3] Among the biased integration methods, steered molecular dynamics (SMD) in combination with the Jarzynski's nonequilibrium work relation[4] has been shown to accurately predict free energy profile of bioprocesses along a predefined steering path such as an unfolding coordinate.

Over the past decade, numerous numerical studies utilizing SMD methodology to obtain the potential of mean force (PMF) along a chosen pulling path have been reported on systems such as stretching of polyalanines,[5–8] unfolding of Ace-Alanine$_8$-NMek,[9] Angeli's salt decomposition,[10] and chicken liver sulfite oxidase modeled in the activated form.[11] The computational costs of these studies, however, are still sufficiently large—especially for larger systems in explicit water—that the number of such studies remains relatively small. One obstacle to the use of SMD is the need for a large number of nonequilibrium trajectories in order to converge the PMF for long distance paths. Many of these trajectories do not contribute to the average because the exponential weights favor only the few paths with the lowest work, thus, lead to inaccurate estimate of the potentials of mean force.[12] The adaptive steered molecular dynamics (ASMD) approach introduced recently by us,[13] overcomes this obstacle by dividing the calculation into smaller segments over which the work distribution is narrower (and often near-Gaussian). In our earlier article,[13] we demonstrated the convergence of this approach for the unraveling of a protein—Neuropeptide Y—in a water solvent, but there was no direct data with which to confirm the specific PMF. Those results did agree indirectly with recent experimental data. The present article demonstrates the efficacy of ASMD through direct comparison to a known system in vacuum, and provides a precise proof of a conjecture in the earlier article.

In Ref. 13, the accuracy of the adaptive SMD methodology—briefly reviewed in Sec. II—could only be inferred from indirect experimental evidence. The critical conjecture underlying ASMD is the selection criterion for restarting a subsequent segment. This could be performed by relaxing the environment while holding the pulling coordinate fixed as suggested by Jarzynski[4] though at considerable computational cost. In Sec. II, we provide a proof that a different selection criterion choosing the configuration from among the nonequilibrium ensemble generated in the prior segment leads to the correct result. In Sec. III A, we demonstrate that ASMD with the selection criterion reproduces the free energy profile of decaalanine stretching obtained earlier in vacuum[5,6] using significantly less CPU time. Therein, we also compare the result to a multi-stage SMD so as to illustrate the relative advantages of ASMD with the selection criterion. In Sec. III B, we also track the number of hydrogen bonds within the peptide along the pulling coordinate of the ASMD so as to illustrate the possibility for obtaining additional observables using ASMD.

a)Author to whom correspondence should be addressed. Electronic mail: hernandez@chemistry.gatech.edu.

Studies[14] of small molecules—e.g., peptides—continue to play a key role as benchmarks of new methods and to test fundamental hypotheses in the context of protein dynamics and folding. Decaalanine, consisting of a mere ten alanine residues, is such an paradigmatic peptide. While necessarily small, the decaalanine peptide is stable in an α-helical structure containing several internal hydrogen bonds that must be broken in order to fully stretch it. It is therefore a good target for the demonstration and verification of methods that probe the energetics of unraveling a peptide in vacuum.[5–7]

II. MODELS AND METHODS

A. Steered molecular dynamics (SMD)

Providing an exact relation between free energy difference and the work done through a directed nonequilibrium process, Jarzynski's equality has been implemented in the context of several experimental[15,16] and computational[17] studies. The process is directed along a preselected path $\lambda(t)$ imposed on a chosen order parameter or reaction path $\xi(\vec{r})$ that marks the dynamical process within the full-dimensional configuration space $\vec{r}$ in some way. In a typical SMD simulation, the system is first driven away from equilibrium by imposing a time-dependent harmonic potential on ξ along the path $\lambda(t)$ through the addition of an auxiliary harmonic potential,

$$h_\lambda(r) = \frac{k}{2}\left(\xi(r) - \lambda\right)^2 \ . \tag{1}$$

Many realizations of the nonequilibrium ensemble are generated. The work $W_{\xi_t \leftarrow \xi_0}$ done by the harmonic restraint to move ξ from its initial ξ_0 to any given value ξ_t along $\lambda(t)$ is calculated for each realization. This work is the so-called accumulated work in the sense of Ref. 18 as was implemented by Jarzynski,[4] Crooks,[17] and Hummer and Szabo.[17] It specifically measures the work applied through the spring from an auxiliary particle that is guided along the specified trajectory $\lambda(t)$. As the system ξ is not rigorously fixed to this point, there is a potential error in the association of the PMF to $\lambda(t)$ as first observed by Hummer and Szabo.[17] Paramore, Ayton, and Voth[19] have shown how to deconvolve the average so as to remove the small error. In the stiff spring limit taken in this work, however, the correction is sufficiently small that this correction can be ignored, and the PMF obtained at $\xi \approx \lambda(t)$ for a given t. Thus the work distribution can be averaged according to Jarzynski's equality

$$G(\xi_t) = G(\xi_0) - \frac{1}{\beta}\ln\langle e^{-\beta W_{\xi_t \leftarrow \xi_0}}\rangle_0 \ , \tag{2}$$

to obtain the PMF along the reaction coordinate.

The exponential averaging of nonequilibrium work—Eq. (2)—implies that the resulting free energy is mostly dominated by the low values of the ensemble work distribution. This suggests that one could define three classes of trajectories: "tight trajectories" whose nonequilibrium work remains low and contributes throughout, "lost trajectories" whose nonequilibrium work becomes so large that they never contribute to the sum once they are lost, and "returning trajectories" that once again contribute to the sum after having been seemingly lost for some period of the nonequilibrium pull. In the limit that only tight trajectories contribute, the SMD calculation reduces to one bounded by the funnel requirement introduced by Wu and Kofke.[20] The implementations of this concept allows one to ignore all trajectories that are lost and hence reduces the computational cost of an SMD calculation considerably.[12,21,22] Specifically, such methods include overlap sampling,[20] Rosenbluth-sampling,[22] and single-ensemble path sampling.[21,23] Wu and Kofke showed the applicability of a two-stage funnel sampling on the nonequilibrium work method.[24] Furthermore, they[20] showed earlier that a multi-stage approach could also be implemented as long as the trajectories connected the intermediate stages through the funnel requirement. However, these leave out the possibility of returning trajectories. The adaptive SMD method discussed here was used by us recently to reduce the required number of trajectories explicitly through the contraction of the nonequilibrium distribution sampled between segments.[13] ASMD shares the spirit of the earlier methods because of the possibility of contracting the number of configurations, but it also permits the possibility of inclusion of returning trajectories not found in the latter. The ASMD includes these returning trajectories implicitly because the resampling of the initial conditions of the solvent at the beginning of each step provides for new trajectories that could have come from non-contributing sojourns in the past. The degree to which such returning trajectories contribute to the average for decaalanine has not, however, been obtained in this article. It should also be noted that the computational advantage of ASMD over some of the previous staging methods lies in the contraction of all the tight trajectories to a common initial structure at the beginning of each stage. This is a more severe contraction than previously employed and substantially reduces the computational cost by avoiding the calculation of many lost trajectories.

B. Adaptive steered molecular dynamics

When a system is in contact with a large enough heat bath, the free energy difference between two equilibrium states can be obtained by the appropriately averaged nonequilibrium work done on the system over an ensemble of transformations between those states as expressed by Eq. (2). In the adaptive SMD algorithm, the overall reaction coordinate ξ is partitioned into N segments bounded by the $N+1$ constrained points, $\xi_0, \xi_1, \ldots, \xi_N$. Within each segment, the work is obtained using Eq. (2) assuming that an appropriate initial configuration is known. The iterated average free energy at any given point ξ_t along the path takes the form,

$$\Delta G_{\xi_t \leftarrow \xi_0} = -\frac{1}{\beta}\sum_{i=1}^{N'} \ln\langle e^{-\beta W_{\xi_i' \leftarrow \xi_{i-1}'}}\rangle_{i-1} \tag{3}$$

for N' chosen such that $\xi_t \in (\xi_{N'-1}, \xi_{N'}]$, $\xi_i' \equiv \xi_i$ for $i < N'$, and $\xi_{N'}' \equiv \xi_t$. The environment and Langevin bath variables (denoted as Γ and Θ, respectively, in Ref. 13) associated with the initial point ξ_0 are initialized by equilibration as in a standard SMD simulation. As discussed below, in the present implementation of ASMD, an ensemble of nonequilibrium

trajectories is obtained within each segment through different random force sequences arising from the bath.

At the end of a given iteration i, the system will be located at ξ_{i+1}, but the environment variables may have been driven quite far from it. Formally, the initial configuration for the subsequent iteration $i + 1$ can be obtained by holding the perturbation fixed—that is ξ held at ξ_{i+1} by an infinitely stiff spring—long enough that the environment relaxes to equilibrium.[5] During this waiting period, no work is done on the system, and hence there is no contribution to Eq. (3). Thus Eq. (3) is a formally exact way of restating Jarzynski's equality in a series of segments. Such an implementation offers little computational advantage in so far as the relaxation stages could be quite expensive. It does, however, offer an advantage in the convergence as the intermediate nonequilibrium trajectories are less likely—because they are shorter—to wander off to distant parts of the landscape. The statistics of the work distribution is consequently more nearly Gaussian, and the convergence of the sum is faster. Echeverria and Amzel[8] essentially followed this procedure in obtaining the helix propensities of dodecaalanine in solvent using a small number (15) of trajectories along a 15 Å stretch.

The computational advantage for ASMD is potentially greater, however, if an efficient criterion can be applied to the selection of the initial configuration at each iteration. Possible choices are the configuration (i) that requires the amount of work that is closest to the Jarzynski's average, (ii) that requires the minimum amount of work, or (iii) that is nearest to the reaction coordinate at the end of the iteration. For simplicity, we will refer to these as the JA, MW, and RC selection criterion, respectively. One objection to the use of such nonequilibrium structures could lie in the possibility that Jarzynski's equality may be applicable only when the SMD is performed between equilibrium structures. However, the applicability of Jarzynski's equality for transformations between nonequilibrium structures has been reported by Hatano and Sasa.[25] Trepagnier *et al.*[26] also suggested that transitions between nonequilibrium steady states are governed by similar laws as transitions between equilibrium states. The resulting potentials of mean force for decaalanine stretching in vacuum obtained using each of the three choices at various stretching rates are shown in Sec. III A. The JA selection criterion—as used in Ref. 13—is seen to be the best of these. This result is not surprising because it amounts to selecting a structure that is fully relaxed—taking advantage of the results from the trajectories that have already been calculated—without requiring an additional relaxation period before initiating the next batch of trajectories.

The validity of the adaptive SMD methodology using the JA selection criterion was illustrated in our earlier work in the case of Neuropeptide Y.[13] We required an ansatz that the process is Markovian and satisfies detailed balance. These are sufficient but not necessary conditions for Jarzynski's nonequilibrium theorems[4] as, for example, Jarzynski's equality has been validated without them.[27,28] Our more-stringent ansatz was also assumed by Schöll-Paschinger and Dellago[29] in extending the nonequilibrium theorems to systems governed by Hamiltonian dynamics, Nosé-Hoover dynamics and Gaussian isokinetic dynamics. It is also related to the funnel requirement of Kofke *et al.*[20] which posits that the accuracy of the free energy calculation is determined by the overlaps in the configurations along the sampled nonequilibrium paths. Specifically, the ansatz is that the ensemble of nonequilibrium configurations at the end of a given iteration is inclusive of the equilibrium ensemble. If so, then one can draw the relaxed structure directly from this ensemble as adopted in ASMD. But which one? As the free energy change $\Delta G_{t\leftarrow t'}[\equiv G(\xi_t) - G(\xi_{t'})]$ of the relaxed structure within a stage with endpoints at t and t' corresponds to the free energy difference according to the change in the PMF dictated by the Jarzynski equality, then the selected structure—with respect to ξ and all other unconstrained internal coordinates—should have arisen from an applied work that is equal (or closest to) the average Jarzynski work. That is, the initial structure for each subsequent iteration is taken to have the same reaction path—ξ—and corresponding phase space variables—Γ—as that for the last point of the trajectory whose work difference was closest to the Jarzynski average. The swarm of nonequilibrium trajectories obtained within each segment are subsequently distributed because the bath variables (Θ) are stochastic and initiated with different random numbers. As long as the average Jarzynski work for the trajectories is dominated by a single basin (moving along ξ_t), then ASMD will give the same result as SMD given sufficient trajectories entering both averages. ASMD, however, requires fewer trajectories because noncontributing trajectories are discarded through the contraction at each iteration. If, instead, there exists multiple disconnected basins given the constraints at ξ_i, then the implementation of ASMD described here with the use of only one branch does not suffice. The existence (or lack thereof) of these basins can be confirmed during the ASMD simulation by measuring the RMSDs between those structures obtained with work functions near to that of the Jarzynski average. If the RMSDs are small (as is the case in the current work on decaalanine), then a singly branched ASMD suffices. Otherwise, it is still possible that a single branch suffices as long as the structures remain accessible through thermal fluctuations, but this requires a more detailed assessment of the structures and their dynamics. In the latter scenario, it is a simple matter to extend ASMD to include such additional branches as initial configurations in subsequent interactions weighted by their relative abundance up to that point. This generalized strategy, however, has not yet been implemented.

C. Simulation parameters

Molecular dynamics simulations have been carried out using NAMD (Ref. 30) with the CHARMM force field.[31] Decaalanine is fully specified in terms of a 104-atom model with the hydrogens included explicitly. Simulations of the peptide in vacuum were carried out at 300 K. The peptide is aligned so that nitrogens at the termini reside on the z-axis and stretched along the z-direction. Note that, in vacuum, the system is not coupled to a solvent which would give rise to an equilibration temperature. The latter is maintained through the imposition of an effective heat bath represented through random and

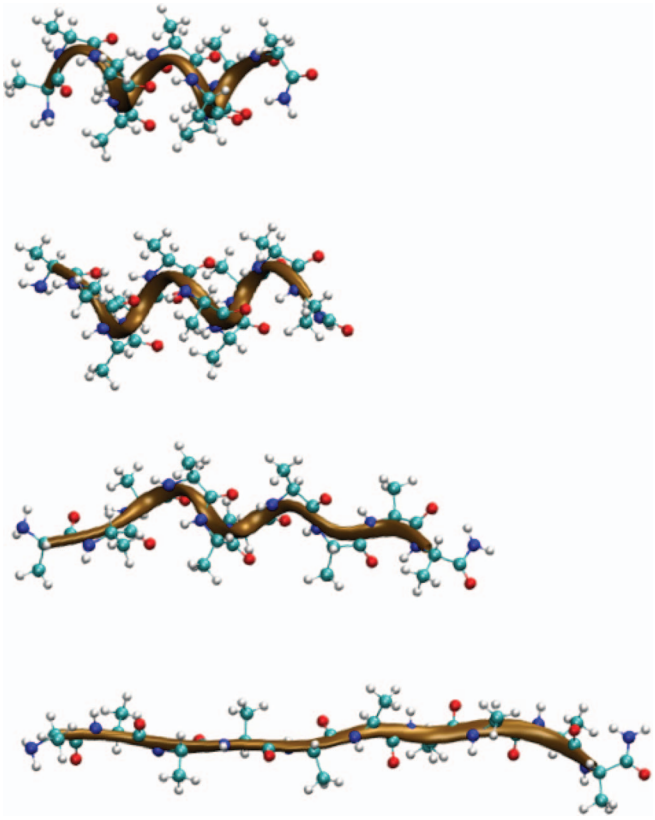

FIG. 1. Representative ribbon and atomically detailed snapshots of decaalanine in vacuum are displayed along the steered path. From top to bottom, the structures correspond to: (a) a compact structure at the NN–NC distance of 13 Å, (b) the minimum energy conformation—an α-helix with an end-to-end distance of 15.2 Å, (c) a structure at the kind of the PMF shown in Fig. 2—at circa 26 Å, and (d) a coil structure at the end of one of the pulled trajectories at the end-to-end distance of 33 Å.

frictional forces—i.e., Langevin dynamics—as was also done by Park and Schulten.[5,6]

The initial coordinates are taken from the compact α-helical structure modeled in vacuum in the earlier SMD studies of Park and Schulten.[5,6] The reaction coordinate is defined as the end-to-end distance between the nitrogen atom of the N-terminus (NN) and the nitrogen atom of the cap at the C-terminus (NC). The stretching of the peptide is imposed using the steering module within NAMD by holding the NN end fixed and directing the NC end relative to the NN end. An ensemble of nonequilibrium trajectories is generated because the thermalizing Langevin bath is initialized differently for each. The overall unraveling coordinate covers the NN–NC distance from 13 Å to 33 Å (Fig. 1). The system is simulated at various stretching speeds in vacuum. The spring constant, $k = 7.2$ kcal/mol, was set fixed to the value used in Refs. 5 and 6. Trajectories are analyzed numerically through metrics chosen according to visual inspection. The latter is facilitated by the NAMD/VMD package.[30,32] PMFs along the decaalanine stretching pathway are calculated for each set of simulations. Hydrogen bonds are identified when donor and acceptor atoms are within 4 Å and the angle formed by donor, hydrogen, and acceptor is greater that 140°.

III. RESULTS AND DISCUSSION

A. The thermodynamics of decaalanine stretching in vacuum

The forced unraveling of decaalanine is investigated in vacuum at two stretching speeds (10 Å/ns and 100 Å/ns). The adaptive SMD simulations are performed in 10 incremental segments as this was found to be sufficient to obtain convergence. This covers a change in the overall stretching coordinate of 20 Å—that is, the NN–NC distance goes from 13 Å $\rightarrow$ 33 Å. For each stretching speed, sets of 50, 100, 200, 400, and 800 trajectories have been simulated at each segment to establish convergence with the number of trajectories. Note that the increase of the number of trajectories requires a nearly complete recalculation of the entire adaptive SMD. This is necessary (in the strictest sense) because the configuration chosen at the end of a given iteration segment and utilized to initialize the next iteration segment is invariably different given the introduction of more trajectories. In standard implementations, however, one could carry out sets of ASMD runs with each set utilizing a different configuration (as chosen from only those trajectories with the given set) leading to better averaging and more efficient parallelization.

The PMFs obtained by ASMD are shown in Fig. 2 according to the JA, MW, and RC criteria. For exact comparison, we have also reproduced the PMFs obtained by Park and Schulten[5,6] within the standard error of the calculations. These are shown as black curves in Fig. 2 and were obtained using 10 000 standard SMD trajectories at a forced speed of 100 Å/ns and 10 Å/ns at left and right, respectively. As per Park and Schulten,[5] the SMD result obtained at the slower pulling speed—10 Å/ns—is also identical to the reversible

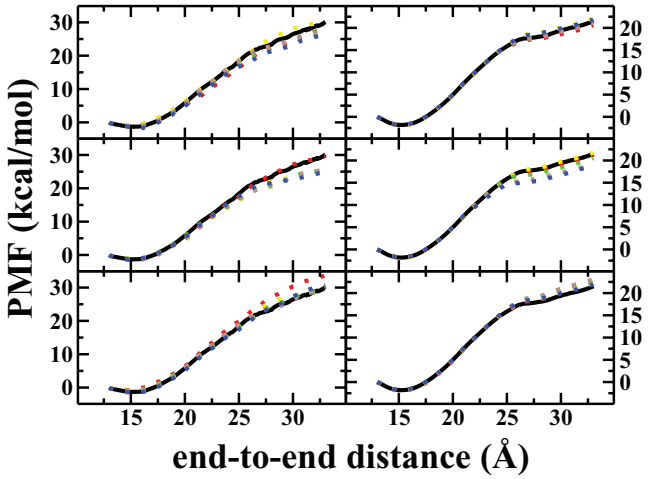

FIG. 2. The comparison of the PMFs obtained from the adaptive SMD method pulling at 100 Å/ns (left column of panels) and 10 Å/ns (right column of panels) when a different selection criterion is used to choose the configuration from the structures at the end of each segment. The configuration is chosen according to the JA, MW, and RC criterion and displayed in the bottom, middle and top panels, respectively. Dashed curves in red, yellow, green, brown and blue represents 50, 100, 200, 400 and 800 trajectories per segment, respectively. The solid black curve is the PMF obtained from averaging 10,000 standard SMD simulations. Note that the standard PMF for the 10 Å/ns pulling simulations (solid black curves in the right column) largely overlaps onto the reversible PMF (although not shown).

PMF calculation. The agreement between ASMD and SMD seen in Fig. 2 is remarkable. Note that ASMD converges to the SMD result at the fast pulling speed, and not to the reversible (and accurate) PMF. This is a consequence of the fact that the error in both is due to the perturbation caused by the pulling speed and not to the nature of the initial distribution. As both are driven out of equilibrium in the same way for a given pulling speed, both converge to the same estimated PMF in said case. Indeed, at both pulling speeds, the ASMD result is equal to the SMD result during the first 10 Å of the end-to-end distance even with a mere 100 trajectories. As will be shown in more detail in Sec. III B, this regime is dominated by the energetics of the breaking and making of internal hydrogen bonds. This evidently steers the nonequilibrium trajectories within a narrow work distribution which can be represented by a small number of trajectories. Beyond 10 Å, convergence is achieved with 800 trajectories. This represents an improvement in performance for ASMD over SMD at a given pulling speed of more than a factor of 10 because the computational cost in both methods is determined by the amount of time needed to integrate trajectories—circa 6000 in SMD and 800 in ASMD—to obtain the converged result. The relative advantages of the JA criterion, the convergence of ASMD, and the contrast to multi-stage SMD including relaxation stages are discussed below in turn.

The JA criterion is the optimal choice for the selection of the initial structure in ASMD. Jarzynski's equality can be used smoothly to obtain the PMF along a series of steered segments through the strong requirement that the trajectories be equilibrated in between segments. The equilibrium distribution at the beginning of each segment can be generated by relaxing the structures at the end of the previous segment through a series of trajectories propagated for a finite time during which no work is done on the system. However, this is potentially cost-prohibitive. Thus, the determination of a suitable criterion for choosing the initial configurations at the beginning of each ASMD segment—without requiring such additional simulations—provides a significant potential savings in computational effort.

The key idea underlying the selection criterion discussed above is that a single representative structure can be selected from the nonequilibrium distribution of structures at the end of a segment to initialize the subsequent segment. Once initialized, the subsequent nonequilibrium steering segment will give rise to the correct distribution as long as: (1) the overall system must be ergodic with fixed λ and (2) the nonequilibrium disturbance must be sufficiently slow so as to allow the system to attain a local equilibrium distribution of structures before any significant work has been done. If these are satisfied, then the system essentially recreates the equilibrium distribution as per the strong requirement for a multi-stage SMD to be applicable. The choice of criterion, however, strongly affects the convergence of the result. To illustrate this point, we have explored three possible selection criteria:

(i) The configuration that requires the amount of work closest to Jarzynski's average.
(ii) The configuration that requires the minimum amount of work.
(iii) The configuration that is the nearest to the reaction coordinate.

The PMFs in Fig. 2 are calculated using a pulling speed of 100 Å/ns (left column) and 10 Å/ns (right column), respectively. In both, the top, middle, and bottom panels correspond to the RC, MS, and JA criterion, respectively.

The choice of MW tends to distort the average towards lower values as the number of trajectories increases for a fixed pulling speed. This is evident in Fig. 2 with the increasing number of trajectories per segment leading to a worse result rather than convergence to the correct result. In principle, a slower pulling speed would lead to a narrower work distribution such that the minimum work trajectory would remain within a fluctuation of the reversible ensemble. Consequently, this should formally lead to a reasonable result given sufficiently slow pulling. However, the numerical results clearly indicate that this would be cost prohibitive.

The choice of the RC leads to averages that oscillate unpredictably around the averaged work. This choice is tantamount to moving the ensemble positions back to the pulling path and is the result of the addition of energy to move the system from a given configuration to that of the RC. As such, the application of the RC constraint effectively does work on the peptide to steer it, but the value of this work is not accounted for. As this work can be positive or negative, it leads to the seemingly random increases and decreases in the PMF. The accuracy should improve, however, with slower pulling speeds and increasing the number of trajectories per step. As the peptide is forcibly stretched at slower speeds, the swarm of trajectories will track the path better, and the magnitude of the unaccounted work will decrease. However, like the MW choice, this convergence will be slow and cost prohibitive.

The choice of JA leads to averages that clearly converge well to the SMD results for a given pulling speed, and in particular to the exact reversible work when the SMD is slow enough. In these cases, it occurs with as little as 400–800 trajectories per segment which is substantially less than the 10 000 trajectories that are required to converge the standard SMD. Thus the adaptive SMD using the JA criterion appears to be both accurate and efficient in terms of CPU requirements. It is also not surprising that it is the only one of the three selection criteria that achieves convergence without requiring a quasi-equilibration trajectory at the outset. As per the proof presented above, the initial trajectory in the JA selection is tantamount to the relaxed structure which would have been obtained with a sufficiently long relaxation stage. However, it is only one such structure and not an ensemble of structures, and hence does not lead to the converged accuracy of a multi-stage SMD including relaxation stages as discussed below.

To better visualize the convergence of the PMFs calculated using ASMD according to the given selection criteria, the relative root-mean-square (RMS) errors are displayed in Fig. 3. The bottom panel exhibits the error relative to the corresponding SMD PMFs for a given pulling speed. The top panel exhibits the error relative to the reversible PMF for all cases. Both comparisons are instructive. The comparison to the corresponding SMD generated PMFs quantifies the

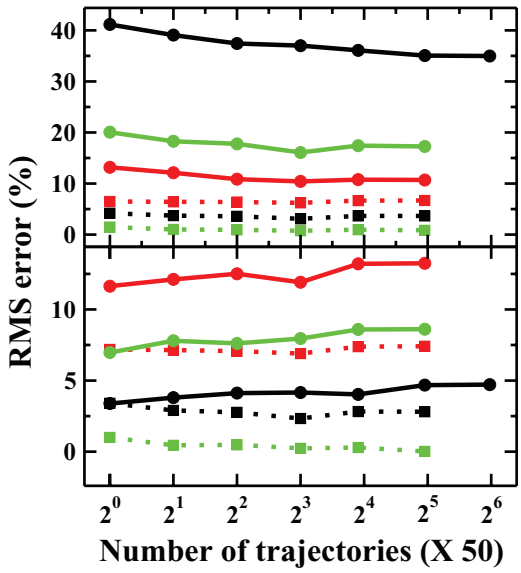

FIG. 3. The convergence of the PMF as a function of the number of sampled trajectories in ASMD (JA, MW, and RC simulations) is shown through the relative root-mean-square (RMS) error in the total free energy difference between the initial and final points. In the top panel, the reference is the reversible PMF for which the total energy difference between the end points is 21.271 kcal/mol. In the bottom panel, the reference is the SMD for which the total energy difference between the end points is 30.134 kcal/mol at 100 Å/ns pulling speed, and 21.516 kcal/mol at Å/ns pulling speed. The abscissa displays the number of trajectories along a range from 50 ($=50 \times 2^0$) to 3 200 ($=50 \times 2^6$). Circle data points connected by the same colored solid lines were obtained from simulation with 100 Å/ns pulling rate; whereas, square data points connected by the same colored dashed lines were obtained from simulation with 10 Å/ns pulling rate. Black, red, green represent JA, MW, RC selection criteria, respectively.

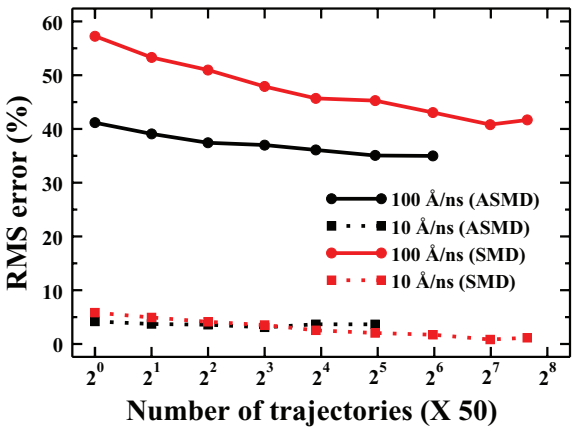

FIG. 4. The convergence of the PMF as a function of the number of sampled trajectories in ASMD (JA selection criterion) and SMD is shown through the relative root-mean-square (RMS) error in the total free energy difference between the initial and final points. As in the top panel of Fig. 3, the reference is the reversible PMF and the number of trajectories are displayed with the same scales in the abscissa. Circle data points connected by the same colored solid lines were obtained from simulation with 100 Å/ns pulling rate; whereas, square data points connected by the same colored dashed lines were obtained from simulation with 10 Å/ns pulling rate. Black and red represent ASMD (JA) and SMD, respectively.

convergence of the ASMD to the SMD result when the JA criterion is employed as discussed above in the context of Fig. 2. Relative to the reversible PMF, however, the MW criterion achieves the lowest RMS errors at the 100 Å/ns pulling speed, whereas the RC criterion achieves the lowest RMS errors at the 10 Å/ns pulling speed. However, the corresponding curves for the MW and RC criteria are not reliable because they are unconverged as seen by the fact that they have not flattened out, whereas they are converged for the JA criterion. The fact that the JA criterion leads to a convergence of the PMFs to the corresponding SMD results thus leads us to conclude that it is the optimal choice for the selection criterion in ASMD.

PMF converges as the number of trajectories is increased. According to the Jarzynski's equality, as the number of the irreversible trajectories increases, the estimated PMF should converge towards the exact PMF obtained from reversible simulations. The catch is that the standard implementation requires many more such trajectories as the end-to-end distance is pulled farther from the original structure. As illustrated in Fig. 2, the adaptive SMD algorithm, however, requires a relatively small number of trajectories—as few as 200 is sufficient to approximately reproduce the PMF for the stretching of decaalanine from a helix to a coil in vacuum. When the JA selection criterion is utilized, the PMF does not differ much as the number of trajectories is increased beyond 200 (bottom rows in Fig. 2). This suggests that for this rather small system—decaalanine in vacuum—200 trajectories per segment is enough to obtain a PMF that would nearly overlap onto the PMF calculated from 10 000 standard SMD trajectories. For larger systems, however, more trajectories may be needed. In the case of the unfolding of Neuropeptide Y (Ref. 13) in water, for example, we found earlier that convergence could be achieved with a few hundred trajectories. Another example is the stretching of decaalanine in explicit water which is being investigated and will be reported separately.[33]

To better visualize the convergence of the PMFs calculated using ASMD, the relative root-mean-square errors are displayed in Fig. 4 with respect to the exact PMF at the end-points. These are compared to the corresponding convergence for the RMS errors obtained using SMD through the same approach taken by Schulten and Park in constructing Figure 6 in Ref. 6. The end-point convergence of the ASMD simulations are obtained using a pool of trajectories obtained for a given pulling speed by accumulating all N_{total} trajectories that had been calculated in the ASMD runs in the convergence plots of Fig. 2. The pool is then randomly divided into non-overlapping sets of a specified number of trajectories, N_{set}. The PMF for each set is obtained, and the corresponding relative end-point RMS error is calculated and reported as a function of N_{set}. In the case of 10 Å/ns pulling, N_{total} was 1550. In the case of 100 Å/ns pulling, N_{total} was 3150 because four additional 400-trajectory ASMD PMFs were generated in order to confirm that the initial conditions did not affect substantially the reported PMFs. When N_{set} exceeds half the size of N_{total} at most one such non-overlapping set can be accommodated. For example, two sets with N_{set} equal to 750 was used to obtain the corresponding RMS rather than a single set of 800. For standard SMD, the pool consists of all 10 000 trajectories sampled as was also implemented by Park and Schulten.[6] For consistency, when possible, the same choices of N_{set} implemented in the ASMD calculation have been used in estimating the convergence of SMD.

As seen in Fig. 4, the relative RMS error of the ASMD calculations in the 100 Å/ns simulations converges at around 36.2% ± 1.2% using as few as 200 trajectories per step. Standard SMD simulations converge at 41.9% ± 1.2% after 3150 trajectories. ASMD outperforms standard SMD in terms of both the percentage of convergence and number of trajectories needed to achieve that convergence. For the 10 Å/ns simulations, ASMD converges at around 3.4% ± 0.3% with only 100 trajectories per step. Standard SMD, at this pulling speed, converges at a lower percentage—i.e., 1.0% ± 0.2%—however, only after approximately 6350 trajectories. While the SMD converges to a slightly better RMS error at 10 Å/ns, the ASMDs need for much fewer trajectories offers a substantial advantage. Thus for both pulling speeds, ASMD displays a more consistent convergence and appears to be less dependent to the number of trajectories required at each step.

As the number of ASMD segments is increased, a larger number of trajectories may be required to obtain convergence. Adaptive SMD is based on dividing the reaction coordinate into certain number of segments so as to apply importance sampling based on the energetics of each trajectory. Specifying an optimum size for the simulation window for each segment is very important in ASMD. One should be careful to choose a large enough segment size for a given pulling speed that the thermalized system is able to spread ergodically across the nonequilibrium ensemble space, but short enough that the work distribution remains near-Gaussian. In this sense, if the segment size is too small (i.e., the reaction coordinate is divided into too many segments), the ensemble reached at the end of each segment will strongly depend on the initial randomization and thus will lead to nonrepresentative sampling of the final distribution. The configuration selected from this non-converged distribution will thereby not represent the initial equilibrium ensemble of the new segment correctly leading to propagation of error. On the other hand, if the bin size is too large (i.e., the reaction coordinate is divided into too few segments) then—depending on how many $k_B T$s of work are applied on the system and the number of trajectories in the ensemble—the work distribution will become non-Gaussian. In this case, the estimated PMF will be dominated by a few trajectories—those rare trajectories with lowest energy—and leading to a poorly converged estimate of the PMF. As shown in Fig. 5, when stretching decaalanine at a speed of 100 Å/ns (top), the PMF is oddly overestimated as the number of segments is increased—note that the PMF from 10-segment-simulations perfectly overlaps onto the PMF from 10 000 standard SMD simulations. When stretching at a lower speed, 10 Å/ns (bottom), no significant deviation is observed. This behavior is expected because the PMF at this (slower) pulling speed yields the exact PMF seen in reversible simulations (when the system is driven away from equilibrium slow enough that the initial spatial and energetic fluctuations are almost negligible). Thus the implementation of ASMD requires convergence both with respect to the number or segments and the number of trajectories employed to obtain the weighted sums.

Segmented SMD connected by relaxation stages does not necessarily give rise to the same PMF as ASMD. The key advance in the ASMD method is that the selection of a single

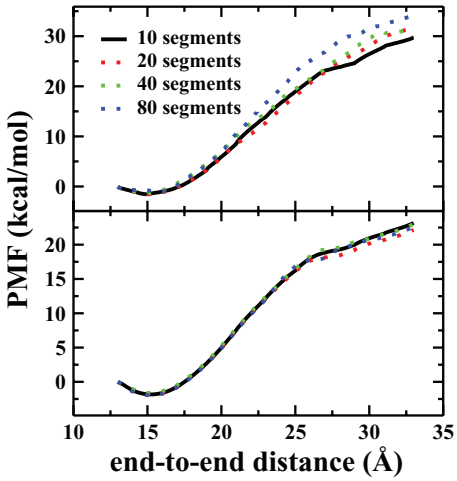

FIG. 5. The comparison of the PMFs obtained from the adaptive SMD method pulling at 100 Å/ns (top) and 10 Å/ns (bottom) when the overall simulation window is divided into 10, 20, 40, and 80 segments.

structure according to the selection criterion gives rise to the same PMF as a multi-stage SMD connected by zero-work relaxation stages. Here we demonstrate that in the decaalanine case, this equivalence may be lost if full relaxation stages are introduced between each segment. During a relaxation stage inserted before the beginning of the ith segment, the reaction coordinate—end-to-end distance—for each trajectory is held fixed at the position reached at the end of $(i-1)$th segment. The remainder of the system is propagated as before for a fixed time increment. This relaxation involves no work on the chosen coordinate and hence does not contribute to the PMF. Specifically, we have repeated the ASMD simulation with 400 trajectories per segment using multi-stage SMD including relaxation stages in between each segment. In five separate ASMD simulations, the system is allowed to propagate for 2 ps, 10 ps, 50 ps, 100 ps, and 200 ps in between SMD segments. Figure 6 displays the comparison of the PMF obtained

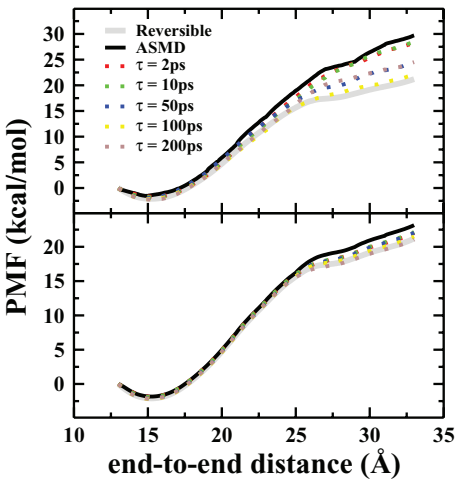

FIG. 6. The PMFs obtained from the adaptive SMD method pulling at 100 Å/ns (top) and 10 Å/ns (bottom) are shown as a function of decaalanine end-to-end distance. The ensemble of trajectories is relaxed at constant temperature and end-to-end distance for 2 ps (red), 100 ps (green), and 200 ps (blue) between pulling segments. The solid black curve is the PMF obtained using ASMD with the JA selection criterion (and no relaxation between segments). The solid grey curve is the reversible PMF.

from these simulations to the ASMD. At both pulling rates, relaxation leads to deviation with the ASMD PMF (cf. Fig. 2) when only short-time relaxation is imposed between steps. (Note that the PMF obtained from ASMD with 400 trajectories per step perfectly overlaps onto the original estimation of the PMF—by Schulten and Park[5]—using 10 000 standard SMD realizations.) Interestingly, the use of constrained relaxation of longer times (50 ps and longer) between the segments leads to a lowering of the calculated PMF towards a better agreement with the exact (and reversible) PMF. It is quite remarkable that this agreement occurs even in the faster 100 Å/ns pulling speed simulations. However, it is not a converged result as an increase in the relaxation time leads to a deviation from the reversible PMF. This key finding suggests that the constrained relaxation has not been propagated sufficiently long to sample the most likely basins. Meanwhile, the ASMD does not suffer from this convergence problem, presumably because this system is dominated by a single basin which is captured by the JA selection criterion. It is likely that the multi-stage SMD including longer relaxation would once again converge to the correct answer by increasing the number of trajectories but this makes it cost-prohibitive.

B. Decaalanine structure and hydrogen bonding along the stretching coordinate

The converged PMF shown in the bottom panel of Fig. 2 reveals important structural properties of the helix-coil transition of decaalanine in vacuum. The initial structure is evidently not the minimum energy structure as the free energy minimum appears at 15.2 Å in nearly all the simulations roughly independent of stretching speed. The initial conformation of decaalanine is a compact helix with unspecified helical character whereas the minimum energy conformations of decaalanine observed at 15.2 Å are a well-defined α-helix. This and other structural properties of peptide along the PMF can be more directly confirmed through a hydrogen bond analysis of the nonequilibrium ASMD ensembles.

Although PMFs were calculated for various ensembles and selection criteria, the hydrogen bond analysis shown in this section is realized over a single criterion (JA—selecting the configuration that requires the amount of work closest to Jarzynski's average) and a single ensemble size (400 trajectories per step). These conditions were seen earlier to be sufficient to obtain a converged PMF; see, for example, the bottom panel of Fig. 2. The sum of hydrogen bonds for a given structure were also partitioned into sums of those bonds linking residues separated by a given number of residues along the chain. For example, the $i \rightarrow i+4$ hydrogen bonds correspond to those bonds between the ith and $(i+4)$th residues of the peptide as would be seen in an α-helix. Averages of the hydrogen bond counts N_{H} are obtained using work averages of the ASMD simulation,

$$\langle N_{\mathrm{H}}(\xi_t)\rangle = \frac{\sum_{i=1}^{N} \hat{N}_{\mathrm{H}}(i,\xi_t) e^{-\beta W_i(\xi_t)}}{\sum_{i=1}^{N} e^{-\beta W_i(\xi_t)}}, \qquad (4)$$

where $N(=400)$ is the number of trajectories and $\hat{N}_{\mathrm{H}}(i,\xi_t)$ is the given hydrogen-bond count for trajectory i at the extension ξ_t of the peptide chain.

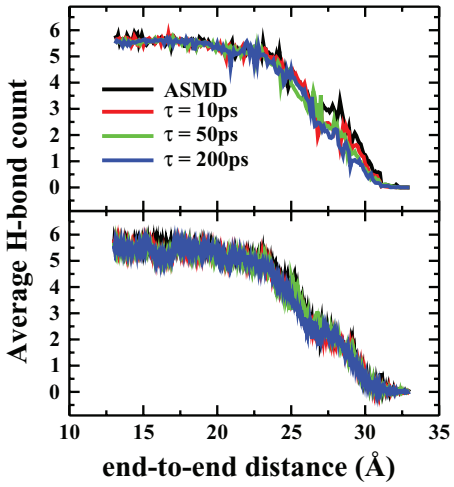

FIG. 7. The average number of internal hydrogen bonds in decaalanine in vacuum is shown as a function of decaalanine end-to-end distance from 100 Å/ns (top) and 10 Å/ns (bottom) pulling simulations. All curves are labeled as in Fig. 6.

In all of the simulations, the average number of intrapeptide hydrogen bonds seen in the initial compact form of the peptide is six and gives rise to its helical character. Stabilizing bonds are maintained with little loss (as shown in Fig. 7) through nearly 10 Å of extension. This is remarkable as it corresponds to nearly half of the final extension at which the peptide is nearly linear. The PMF over this region (cf., Fig. 2) indicates a large increase in the free energy with increasing extension. After this point, the number of helix-stabilizing hydrogen bonds gradually decrease up to 31 Å separation. Meanwhile, the PMF over this region (cf., Fig. 2) indicates a smaller increase in the free energy with increasing extension. At this point no more helical contacts are present and the peptide is an extended random coil. The puzzle, answered below, is what are the origins of the dramatic free energy change with little change in H-bonding in the first half of the extension while the converse occurs in the second half of the extension.

Figure 7 also shows a comparison of the hydrogen bond count between the ASMD and the multi-stage SMD (including relaxation stages in between) as described earlier. In agreement with the PMF analysis (cf., Fig. 6), intrapeptide hydrogen bond breakage and reformation appears to be similar in the two approaches. (Note that averages from $\tau = 2$ ps and 100 ps are not shown for clarity; they also overlapped with the others.) In fact, as seen in the bottom panel of Fig. 7, from 10 Å/ns pulling simulations—which estimates the reversible energetics almost perfectly—no clear distinction is observed in the hydrogen bond analysis. Any discontinuity observed in the hydrogen bond count is almost instantly shifted towards the accurate count (see blue curve that represents 200 ps relaxation at the beginning of each segment). When faster pulling is implemented (top panel of Fig. 7, however, the discontinuities are not easily shifted. Longer relaxation—50 ps and 200 ps—slightly enhances this problem as seen in the green and blue data in Fig. 7, respectively. However, the discontinuities do not disappear completely. This is mainly due to the hydrogen bond forming affinity of the nitrogen at the N-terminal. The change in the intrapeptide hydrogen bond behavior is rather

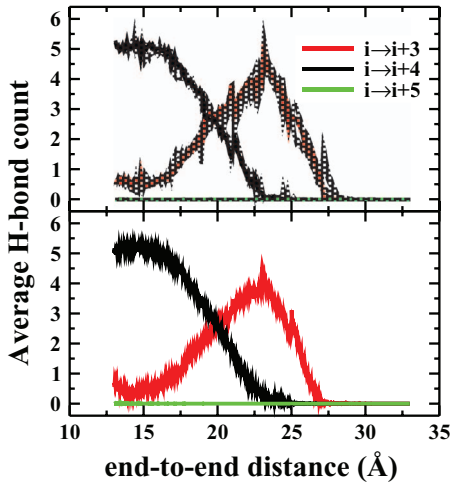

FIG. 8. The average number of internal hydrogen bonds in decaalanine as a function of decaalanine end-to-end distance is shown for fast pulling (top panel) and slow pulling (bottom panel) simulations. Black represents $i \rightarrow i + 4$ (α-helix), red represents $i \rightarrow i + 3$ (3_{10}-helix), and green represents $i \rightarrow i + 5$ (π-helix). The semi-transparent curves in the top panel correspond to five additional independent simulations and indicate the spread in the error. Each of these gave rise to a PMF which is the same as that shown earlier within the resolution of the plots.

limited because the two nitrogen atoms at the termini are held fixed during relaxation. This behavior is again in agreement to the energetic fluctuations seen in Fig. 6. The top panel of Fig. 7 displays significant discontinuity in the hydrogen bond counts especially when the relaxation time is limited to below 50 ps.

In order to resolve the puzzle associated with the apparent disconnect between the trends in the total hydrogen-bonding and the PMF, the hydrogen-bonds were partitioned according to the chain distance, and the averages are shown in Fig. 8. In the ASMD simulations of the decaalanine stretch, the α-helical contacts ($i \rightarrow i + 4$) break within a few Angstroms and not the 10 Å window during which the total number of hydrogen bonds remains fixed near six. Indeed, it appears that the α-helical contacts are mostly and immediately replaced by the 3_{10}-helical contacts ($i \rightarrow i + 3$) during the first half of the overall stretch and thus the total remains nearly constant in Fig. 7. The interconversion between α-helix and 3_{10}-helix hydrogen bonds was previously reported to be a key mechanism of the helix-coil transformation.[34] This has been confirmed in this study with the observed replacement of $i \rightarrow i + 4$ hydrogen bonds by $i \rightarrow i + 3$ hydrogen bonds as the decaalanine helix unwinds during the ASMD simulations. When the end-to-end separation reaches 18 Å, there are an equal number of α-helix and 3_{10}-helix hydrogen bonds present in the ensemble, but the diminution of the three hydrogen bonds from $i \rightarrow i + 4$ to $i \rightarrow i + 3$ contacts accounts for the approximately 6 kcal/mol increase in the PMF seen in Fig. 2). In both slow and fast pulling simulations, all α-helical contacts are broken at around 23 Å separation. Not surprisingly, this point is also in near coincidence with the maximum in the number of the 3_{10}-helical contacts. That maximum is near 4 and does not quite entirely replace the 6 broken $i \rightarrow i + 4$ hydrogen bonds lost from the initial helix. The loss of two such hydrogen bonds and the diminution of the other hydrogen bonds from $i \rightarrow i + 4$ to $i \rightarrow i + 3$ contacts, roughly accounts for the 15 kcal/mol increase seen in the PMF seen in Fig. 2). Meanwhile the 3_{10}-helical contacts are not as stable. Consequently, the loss of these weak hydrogen bonds during the second half of the overall stretch gives rise to a small free energy increase as seen in the PMF while there is a dramatic loss of total hydrogen bonds. A careful reader will likely note that the sum of the three helical contacts displayed in Fig. 8 does not necessarily equal the total number of hydrogen bonds displayed in Fig. 7. The difference arises from other possible contacts—e.g., $i \rightarrow i + 2$ or $i + 6$—that have not been reported explicitly herein.

The relative populations of α-helix ($\phi \approx -57°$, $\psi \approx -47°$), 3_{10}-helix ($\phi \approx -49°$, $\psi \approx -27°$), π-helix ($\phi \approx -57°$, $\psi \approx -70°$), and extended-coil ($\phi \approx -90°$, $\psi \approx 120°$) structures can be seen in the top and bottom time-series of Ramachandran plots shown in Fig. 9 for the 100 Å/ns pulling simulations and the 10 Å/ns pulling simulations, respectively. The increase in the number of 3_{10}-helical contacts in the first 1 Å of stretching (i.e., up to the structure when the end-to-end distance is 14 Å) in the vacuum is visible in the broadening of the distribution towards larger ψ. This behavior is followed by the straightening of the decaalanine backbone towards the absolute value of $\psi(\geq 120°)$. As observed in Fig. 8, negligibly few configurations make $i \rightarrow i + 5$ contacts at the end of the first step. This is visible in Fig. 9 as it shows almost zero population of π-helix with $\phi \approx -57°$ and $\psi \approx -70°$. It is notable that the distributions of angles

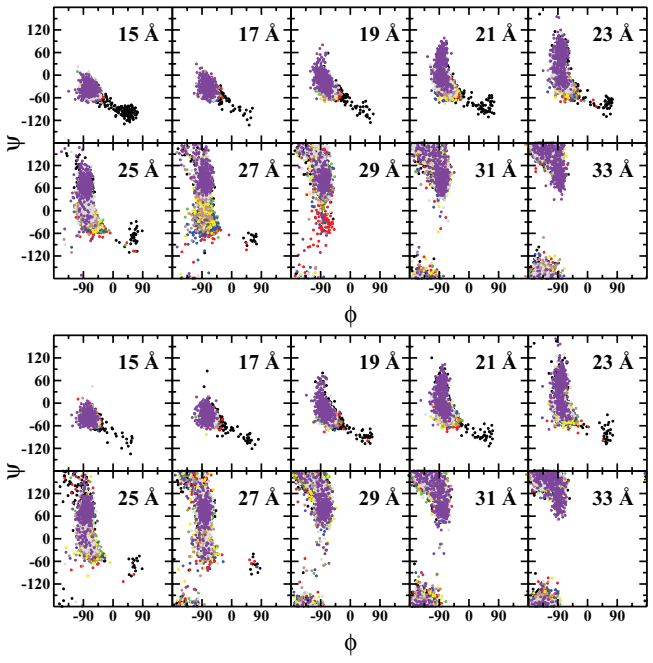

FIG. 9. Ramachandran plots of the middle eight residues (excluding the termini residues because they do not have a pair of ϕ and ψ angles) is displayed for the 100 Å/ns stretching simulations (top) for the 10 Å/ns stretching simulations (bottom) (the selection criterion is JA, the number of trajectories sampled per step is 400 for each of them). Each diagram has a total of 3200 data points: 8 ϕ-ψ angle pairs for each of the 400 trajectories. The reaction coordinate begins from top left and goes towards bottom right by walking along each row. Coloring is as follows: ALA2, black; ALA3, red; ALA4, green; ALA5, blue; ALA6, yellow; ALA7, brown; ALA8, gray; ALA9, purple.

associated with each of the angles are localized over small domains, and that these domains are narrower for the slower SMD pulls. Such behavior is expected if the nonequilibrium trajectories span a single basin of structures, and is therefore consistent with our use of singly branched ASMD in which only one representative structure is selected for initializing each segment according to the selection criterion.

IV. CONCLUSIONS

A recently formulated adaptive SMD (ASMD) method[13] has been used to reproduce the potential of mean force for the stretching of decaalanine in vacuum. Calculated free energy profiles are in agreement with the earlier work of Park and Schulten[5,6] and serve as a confirmation of the ASMD approach. The ASMD method is found to converge to the correct PMF with significantly fewer trajectories as it effectively implements importance sampling over the most probable nonequilibrium configurations. This is achieved by partitioning the reaction coordinate into a series of segments with each initialized according to a selection criterion—namely how the initial structure is somehow chosen from nonequilibrium distribution obtained at the end of the previous segment. In principle if a sufficient number of trajectories are sampled and the steering speed is sufficiently slow, the ASMD PMF should not depend on the selection criterion. Within the maximum number of trajectories per segment sampled in this study, selecting the configuration that required the amount of work closest to the Jarzynski's exponential estimate is found to be the most robust selection criterion numerically. This is consistent with the proof presented in this article that shows that the choice of this energy relaxed structure is equivalent to selecting a structure from a zero-work relaxation-to-equilibrium segment.

If only one branch of nonequilibrium trajectories (accessible to each other through thermal fluctuations at a given constrained point along the steering path) contribute to the nonequilibrium work relation, then the use of a single structure to initialize a given segment is sufficient to achieve convergence using ASMD or multi-stage SMD. Otherwise, representative structures from each branch—as reached through the nonequilibrium trajectories along a different segment—should be used to obtain convergence in the nonequilibrium work. The implementation and verification of multi-branch ASMD will be the subject of future work.

It should also be noted that the implementation of the ASMD strategy may be relevant to the experimental determination of the PMF along a stretching or pulled coordinate. After Hummer and Szabo[17] derived Jarzynski's equality specifically for single molecule pulling experiments such as atomic force microscopy and optical tweezers, these methods have been extensively used to calculate the overall free energy change of the system using the measured force.[15,19,35] Hummer and Szabo later showed that the system free energy can be related to the molecular free energy surface.[36] In a recent examination of Jarzynski's equality using a harmonic spring model, Zimanyi and Silbey[17] showed that the resulting free energy surface does not depend on arbitrary terms in the Hamiltonian—whether it is time dependent or not—and reaffirmed the applicability of the Jarzynski's equality in single molecule force pulling experiments such as atomic force microscopy and optical tweezers, and simulations such as steered molecular dynamics. The selection criterion implemented in the numerical ASMD method is clearly not possible within the experimental system because the initial choice could not be replicated for multiple subsequent pulls. However, the use of alternating short pulls with constrained relaxation stages should allow for much more efficient convergence of the experimental nonequilibrium work functions. The key would be to ensure that the constraint is imposed either through no additional work—as in ASMD—or through a measurable amount of work that is also included in the nonequilibrium averaging.

ACKNOWLEDGMENTS

This work has been partially supported by the National Science Foundation (NSF) through Grant No. CHE 1112067. The computing resources necessary for this research were provided in part by the National Science Foundation through TeraGrid resources provided by the Purdue Dell PowerEdge Linux Cluster (Steele) under grant number TG-CTS090079, and by the Center for Computational Molecular Science & Technology through Grant No. CHE 0946869.

[1] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[2] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[3] R. W. Zwanzig, J. Chem. Phys. **22**, 1420 (1954); T. Straatsma and A. McCammon, Annu. Rev. Phys. Chem. **43**, 407 (1992).
[4] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997); Phys. Rev. Lett. **78**, 2690 (1997).
[5] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, J. Chem. Phys. **119**, 3559 (2003).
[6] S. Park and K. Schulten, J. Chem. Phys. **120**, 5946 (2004).
[7] D. D. L. Minh and J. A. McCammon, J. Phys. Chem. B **112**, 5892 (2008).
[8] I. Echeverria and L. M. Amzel, Proteins: Struct. Funct. Bioinfom. **78**, 1302 (2010).
[9] H. Xiong, A. Crespo, M. Marti, D. Estrin, and A. E. Roitberg, Theor. Chem. Acta **116**, 338 (2006).
[10] J. Torras, G. de M. Seabra, and A. E. Roitberg, J. Chem. Theory Comput. **5**, 37 (2009).
[11] T. Utesch and M. A. Mroginski, J. Phys. Chem. Lett. **1**, 2159 (2010).
[12] M. A. Cuendet and O. Michielin, Biophys. J. **95**, 3575–3590 (2008).
[13] G. Ozer, E. Valeev, S. Quirk, and R. Hernandez, J. Chem. Theory Comput. **6**, 3026 (2010).
[14] W. A. Eaton, V. Muñoz, P. A. Thompson, E. R. Henry, and J. Hofrichter, Acc. Chem. Res. **31**, 745 (1998); V. S. Pande and D. S. Rokhsar, Proc. Natl. Acad. Sci. U.S.A. **96**, 9062 (1999); A. R. Dinner, T. Lazaridis, and M. Karplus, *ibid.* **96**, 9068 (1999); J. C. Crane, E. K. Koepf, J. W. Kelly, and M. Gruebele, J. Mol. Biol. **298**, 283 (2000); T. Beke-Somfai and A. Perczel, J. Phys. Chem. Lett. **1**, 1341 (2010); M. Spichty, M. Cecchini, and M. Karplus, *ibid.* **1**, 1922 (2010).
[15] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Science **296**, 1832 (2002).
[16] F. Douarche, S. Ciliberto, A. Petrosyan, and L. Rabbiosi, Europhys. Lett. **70**, 593 (2005); R. Berkovich, J. Klafter, and M. Urbakh, J. Phys.: Condens. Matter **20**, 354008 (2008).
[17] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998); G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001); D. Rodriguez-Gomez, E. Darve, and A. Pohorille, J. Chem. Phys. **120**, 3563 (2004); P. L. Geissler and C. Dellago, J. Phys. Chem. B **108**, 6667 (2004); F. M. Ytreberg and D. M. Zuckerman, J. Chem. Phys. **120**, 10876 (2004); T. Hatano and S. I. Sasa, Phys. Rev. Lett. **86**, 3463 (2001); D. Wu and D. A. Kofke, J. Chem. Phys. **121**, 8742 (2004); E. N. Zimanyi and R. J. Silbey, *ibid.* **130**, 171102 (2009).

[18]J. M. Schurr and B. S. Fujimoto, J. Phys. Chem. B **107**, 14007 (2003).
[19]S. Paramore, G. S. Ayton, and G. A. Voth, J. Chem. Phys. **126**, 051102 (2007).
[20]N. Lu, D. Wu, T. B. Woolf, and D. A. Kofke, Phys. Rev. E **69**, 057702 (2004).
[21]F. M. Ytreberga and D. M. Zuckerman, J. Chem. Phys. **120**, 10876 (2004).
[22]D. Wu and D. A. Kofke, J. Chem. Phys. **122**, 204104 (2005).
[23]F. M. Ytreberga, R. H. Swendsen, and D. M. Zuckerman, J. Chem. Phys. **125**, 184114 (2006).
[24]D. Wu and D. A. Kofke, J. Chem. Phys. **123**, 084109 (2005).
[25]T. Hatano, Phys. Rev. E **60**, 5017 (1999); T. Hatano and S.-I. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).
[26]E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, Proc. Natl. Acad. Sci. U.S.A. **101**, 15038 (2004).
[27]M. A. Cuendet, Phys. Rev. Lett. **96**, 120602 (2006).
[28]M. A. Cuendet, J. Chem. Phys. **125**, 144109 (2006).
[29]E. Scholl-Paschinger and C. Dellago, J. Chem. Phys. **125**, 054105 (2006).
[30]J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, J. Comput. Chem. **28**, 1781 (2005).
[31]B. R. Brooks, R. E. Bruccoleri, R. E. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[32]W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[33]G. Ozer, S. Quirk, and R. Hernandez, "Adaptive steered molecular dynamics: energetics and hydrogen-bonding of decaalanine stretching in water" (unpublished).
[34]P. Doruker and I. Bahar, Biophys. J. **72**, 2445 (1997).
[35]D. Collin, F. Ritort, C. Jarzynski, J. S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Nature (London) **437**, 231 (2005); R. Nome, J. Zhao, W. Hoff, and N. Scherer, Proc. Natl. Acad. Sci. U.S.A. **104**, 20799 (2007); J. Preiner, H. Janovjak, C. Rankl, H. Knaus, D. Cisneros, A. Kedrov, F. Kienberger, D. Muller, and P. Hinterdorfer, Biophys. J. **93**, 930 (2007).
[36]G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **107**, 21441–21446 (2010).